On Eigenvalue Variations of Rayleigh Quotient Matrix Pencils of a Definite Pencil

Ren-Cang Li
Department of Mathematics
University of California at Berkeley
Berkeley, California 94720

Dedicated to Chandler Davis

Submitted by Shmuel Friedland

ABSTRACT

Let $A - \lambda B$ be a definite matrix pencil of order n, i.e., both A and B are $n \times n$ Hermitian and $c(A, B) \stackrel{\text{def}}{=} \min_{x \in \mathbb{C}^n, \|x\|_2 = 1} |x^H(A + iB)x| > 0$. Suppose Y is an $n \times l$ matrix with full column rank whose column vectors span an approximate invariant subspace for $A - \lambda B$. This note investigates the relation between eigenvalues of $A - \lambda B$ and those of $Y^HAY - \lambda Y^HBY$. Our result for the spectral norm improves Sun's [*Linear Algebra Appl.* 139:253–267 (1990)]. We also present bounds in Frobenius norms and in general unitarily invariant norms.

1. INTRODUCTION

Let A be Hermitian with an eigenvalue λ and the corresponding eigenvector x. The following is well known: For any vector $\tilde{x}$ such that $\sin\theta(x, \tilde{x}) = O(\epsilon)$, we have

$$\tilde{x}^H A \tilde{x} / \tilde{x}^H \tilde{x} = \lambda + O(\epsilon^2)$$

(see, e.g., Parlett [7]). Here the superscript H denotes conjugate transpose. This result has been generalized to the Rayleigh quotient matrix case by Li [2], Liu and Xu [6], and Sun [12]. Li [2] proved an open question raised by Sun [12] and improved Liu and Xu's results.

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

In another direction, the result was generalized to the generalized eigenvalue problems for definite pencils by Sun [11]. In this note, we also consider this kind of generalizations. We improve Sun's result and present new bounds in Frobenius norm and in general unitarily invariant norms.

In what follows, we use $\mathbb{C}^{m\times n}$ for the set of $m \times n$ complex matrices, $\mathbb{U}_n \in \mathbb{C}^{n\times n}$ for the set of $n \times n$ unitary matrices, and I for the identity matrix with suitable dimension which should be clear from the context. We use $\|\cdot\|_2$ for the spectral norm and $\|\cdot\|_F$ for the Frobenius norm. Given $Y \in \mathbb{C}^{n\times l}$, $\mathscr{Y} \stackrel{\text{def}}{=} \mathscr{R}(Y)$ denotes the subspace spanned by the column vectors of Y.

2. PRELIMINARIES

Let $A, B \in \mathbb{C}^{n\times n}$ be Hermitian. $A - \lambda B$ is termed a *definite matrix pencil of order n* if

$$c(A, B) \stackrel{\text{def}}{=} \min_{x\in\mathbb{C}^n,\ \|x\|_2=1} |x^H(A + iB)x| > 0. \tag{2.1}$$

Hereafter, we will use $\mathbb{D}(n)$ for the set of all $n \times n$ definite pencils. A number pair $(\alpha, \beta) \neq (0,0)$ is called a *generalized eigenvalue* of $A - \lambda B$ if $\det(\beta A - \alpha B) = 0$, $0 \neq x \in \mathbb{C}^n$ is an *eigenvector* associated with (α, β) if $\beta Ax = \alpha Bx$. The set of all generalized eigenvalues of $A - \lambda B$ is denoted by $\lambda(A, B)$. There are a lot of equivalent definitions of an eigenspace for $A - \lambda B \in \mathbb{D}(n)$. One of them is: A subspace $\mathscr{Y} \subset \mathbb{C}^n$ is an *eigenspace* of $A - \lambda B \in \mathbb{D}(n)$ if

$$\dim(A\mathscr{Y} + B\mathscr{Y}) \leqslant \dim \mathscr{Y}.$$

The reader is referred to Sun [13] for others. Given $\mathscr{X}_1 \subset \mathbb{C}^n$, an l-dimensional eigenspace of $A - \lambda B \in \mathbb{D}(n)$. Let $X_1 \in \mathbb{C}^{n\times l}$ and $\mathscr{X}_1 = \mathscr{R}(X_1)$. It is known that

$$\lambda(X_1^H A X_1, X_1^H B X_1) \subset \lambda(A, B). \tag{2.2}$$

However, if $\tilde{X}_1 \in \mathbb{C}^{n\times l}$ with rank $\tilde{X}_1 = l$ for which $\mathscr{R}(\tilde{X}_1)$ is close to $\mathscr{X}_1$ in some sense, we cannot expect that (2.2) with X_1 replaced by $\tilde{X}_1$ remains true. It is the purpose of this note to investigate relations between $\lambda(\tilde{X}_1^H A\tilde{X}_1, \tilde{X}_1^H B\tilde{X}_1)$ and $\lambda(A, B)$.

Let $A - \lambda B \in \mathbb{D}(n)$, $Y \in \mathbb{C}^{n\times l}$, and $\operatorname{rank} Y = l$. Define $Y^H AY - \lambda Y^H BY \in \mathbb{D}(l)$ to be the *Rayleigh quotient matrix pencil* of $A - \lambda B$ with respect to Y.

For two nonzero number pairs (α, β) and $(\tilde{\alpha}, \tilde{\beta})$, the *chordal distance* will be used throughout:

$$\rho\big((\alpha,\beta),(\tilde{\alpha},\tilde{\beta})\big) \stackrel{\text{def}}{=} \frac{|\alpha\tilde{\beta} - \beta\tilde{\alpha}|}{\sqrt{|\alpha|^2 + |\beta|^2}\sqrt{|\tilde{\alpha}|^2 + |\tilde{\beta}|^2}}.$$

For two l-dimensional subspaces $\mathscr{X}_1 = \mathscr{R}(X_1)$ and $\tilde{\mathscr{X}}_1 = \mathscr{R}(\tilde{X}_1)$, the distance between them will be measured by $\|\sin\Theta(X_1, \tilde{X}_1)\|$ where $\|\cdot\|$ is a matrix norm and

$$\Theta(X_1, \tilde{X}_1) \stackrel{\text{def}}{=} \arccos\left(\tilde{X}_{10}^H X_{10} X_{10}^H \tilde{X}_{10}\right)^{1/2} \geqslant 0,$$

where $X_{10} = X_1(X_1^H X_1)^{-1/2}$, $\tilde{X}_{10} = \tilde{X}_1(\tilde{X}_1^H \tilde{X}_1)^{-1/2}$. It has been proved that if

$$(X_1, X_2)^{-1} = \begin{pmatrix} W_1^H \\ W_2^H \end{pmatrix}, \qquad \text{where } X_2, W_2 \in \mathbb{C}^{n\times(n-l)} \text{ and } W_1 \in \mathbb{C}^{n\times l},$$

then

$$\rho_p(\mathscr{X}_1, \tilde{\mathscr{X}}_1) \stackrel{\text{def}}{=} \|\sin\Theta(X_1, \tilde{X}_1)\|_p = \left\|(W_2^H W_2)^{-1/2} W_2^H \tilde{X}_{10}^H\right\|_p \qquad (2.3)$$

for $p = 2, F$ (see, e.g., Li [2]).

For definite pencils, we have the following fundamental result due to Stewart [8].

LEMMA 2.1. *Let $A - \lambda B \in \mathbb{D}(n)$. Then there is a nonsingular matrix $X \in \mathbb{C}^{n\times n}$ such that*

$$X^H AX = \operatorname{diag}(\alpha_1, \ldots, \alpha_n), \qquad X^H BX = \operatorname{diag}(\beta_1, \ldots, \beta_n).$$

It is easy to see that α_i and β_j are all real, and can be made such that $\alpha_j^2 + \beta_j^2 = 1$. The following result is due to [9].

LEMMA 2.2. *In Lemma* 2.1, *if* $\alpha_j^2 + \beta_j^2 = 1$ *for* $1 \leqslant j \leqslant n$, *then*

$$\|X\|_2 \leqslant \frac{1}{\sqrt{c(A,B)}}, \qquad \|X^{-1}\|_2 \leqslant \frac{\|(A,B)\|_2}{\sqrt{c(A,B)}}. \tag{2.4}$$

For eigenvalue variations, Sun [10] proved the following:

LEMMA 2.3. *Let* $A - \lambda B \in \mathbb{D}(n)$, *and let* $\tilde{A} = A + E$, $\tilde{B} = B + F$ *be* $n \times n$ *Hermitian matrices. If*

$$\max_{\|x\|_2=1} \sqrt{\frac{(x^H E x)^2 + (x^H F x)^2}{(x^H A x)^2 + (x^H B x)^2}} < 1,$$

then $\tilde{A} - \lambda \tilde{B} \in \mathbb{D}(n)$. *Moreover, let* $\lambda(A, B) = \{(\alpha_i, \beta_i),\ i = 1, \dots, n\}$ *and* $\lambda(\tilde{A}, \tilde{B}) = \{(\tilde{\alpha}_j, \tilde{\beta}_j),\ j = 1, \dots, n\}$. *There exists a permutation* τ *of* $\{1, \dots, n\}$ *such that*

$$\max_{1 \leqslant j \leqslant n} \rho\Big((\alpha_j, \beta_j), \big(\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}\big)\Big) \leqslant \max_{\|x\|_2 = 1} \rho\Big((x^H A x, x^H B x), (x^H \tilde{A} x, x^H \tilde{B} x)\Big).$$

3. MAIN RESULTS

THEOREM 3.1. *Let* $A - \lambda B \in \mathbb{D}(n)$ *with* $\lambda(A, B) = \{(\alpha_j, \beta_j),\ j = 1, \dots, n\}$, *and let* $\mathscr{X}_1 = \mathscr{R}(X_1)$ *be the eigenspace of* $A - \lambda B$ *associated with* $\{(\alpha_j, \beta_j),\ j = 1, \dots, l\}$, *where* $X_1 \in \mathbb{C}^{n \times l}$. *Assume* $\tilde{\mathscr{X}}_1 = \mathscr{R}(\tilde{X}_1)$ *is an l-dimensional approximate eigenspace of* $A - \lambda B$ *such that*

$$\eta \stackrel{\text{def}}{=} \frac{\|(A,B)\|_2}{c(A,B)} \rho_2\big(\mathscr{X}_1, \tilde{\mathscr{X}}_1\big) < 1. \tag{3.1}$$

Let $\lambda(\tilde{X}_1^H A \tilde{X}_1, \tilde{X}_1^H B \tilde{X}_1) = \{(\tilde{\alpha}_j, \tilde{\beta}_j),\ j = 1, \dots, l\}$. *Then there is a permutation* τ *of* $\{1, \dots, l\}$ *such that*

$$\max_{1 \leqslant j \leqslant l} \rho\Big((\alpha_j, \beta_j), \big(\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}\big)\Big) \leqslant \max_{\substack{1 \leqslant i \leqslant l \\ l+1 \leqslant j \leqslant n}} \rho\big((\alpha_i, \beta_i), (\alpha_j, \beta_j)\big) \cdot \eta^2. \tag{3.2}$$

Theorem 3.1 improves Sun's theorem [11] in two respects:

(1) Instead of (3.1), Sun assumes

$$\frac{\max\left\{\sqrt{\|(A_1, B_1)\|_2}, 1\right\}}{\min\left\{\sqrt{\lambda_{\min}(A_1^2 + B_1^2)}, 1\right\}} \cdot \frac{\|(A, B)\|_2}{c(A, B)} \rho_2(\mathscr{X}_1, \tilde{\mathscr{X}}_1) < 1,$$

which is stronger than our assumption (3.1). Here $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of a Hermitian matrix; $A_1 = X_{10}^H A X_{10}$, $B_1 = X_{10}^H B X_{10}$, and $X_{10} = X_1(X_1^H X_1)^{-1/2}$.

(2) The bound in (3.2) improves Sun's bound by a factor $\|(A, B)\|_2 / c(A, B)$.

Proof of Theorem 3.1. It follows from $\mathscr{X}_1$ being an l-dimensional eigenspace of $A - \lambda B$ that there is $X_{20} \in \mathbb{C}^{n\times(n-l)}$ with $X_{20}^H X_{20} = I$ such that

$$X_0^H A X_0 = \begin{pmatrix} A_1 & \\ & A_2 \end{pmatrix}, \qquad X_0^H B X_0 = \begin{pmatrix} B_1 & \\ & B_2 \end{pmatrix}, \tag{3.3}$$

where $X_0 = (X_{10}, X_{20}) \in \mathbb{C}^{n\times n}$ nonsingular, $X_{10} = X_1(X_1^H X_1)^{-1/2}$, and $A_1 - \lambda B_1 \in \mathbb{D}(l)$, $A_2 - \lambda B_2 \in \mathbb{D}(n-l)$. By Lemma 2.1, there are $P_1 \in \mathbb{C}^{l\times l}$, $P_2 \in \mathbb{C}^{(n-l)\times(n-l)}$, both nonsingular, such that for $i = 1, 2$

$$P_i^H A_i P_i = \Lambda_i, \qquad P_i^H B_i P_i = \Omega_i,$$

where

$$\Lambda_1 = \operatorname{diag}(\alpha_1, \dots, \alpha_l), \qquad \Omega_1 = \operatorname{diag}(\beta_1, \dots, \beta_l),$$

$$\Lambda_2 = \operatorname{diag}(\alpha_{l+1}, \dots, \alpha_n), \quad \Omega_2 = \operatorname{diag}(\beta_{l+1}, \dots, \beta_n),$$

and $\alpha_j^2 + \beta_j^2 = 1$, $j = 1, \dots, n$. Set $X_i = X_{10} P_i$ for $i = 1, 2$ and $X = (X_1, X_2)$. Then we have

$$X^H A X = \operatorname{diag}(\Lambda_1, \Lambda_2), \qquad X^H B X = \operatorname{diag}(\Omega_1, \Omega_2).$$

Partition X^{-1} as

$$X^{-1} = \begin{pmatrix} W_1^H \\ W_2^H \end{pmatrix} \qquad \text{with} \quad W_1 \in \mathbb{C}^{n\times l}.$$

Now we get

$$\begin{aligned} \tilde{X}_{10}^H A\tilde{X}_{10} &= \tilde{X}_{10}^H X^{-H} \begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix} X^{-1}\tilde{X}_{10} \\ &= \left(W_1^H\tilde{X}_{10}\right)^H \Lambda_1 W_1^H \tilde{X}_{10} + \left(W_2^H\tilde{X}_{10}\right)^H \Lambda_2 W_2^H \tilde{X}_{10} \\ &\stackrel{\text{def}}{=} \tilde{A}_1 + E. \end{aligned} \tag{3.4}$$

Similarly, we have

$$\begin{aligned} \tilde{X}_{10}^H B\tilde{X}_{10} &= \left(W_1^H\tilde{X}_{10}\right)^H \Omega_1 W_1^H \tilde{X}_{10} + \left(W_2^H\tilde{X}_{10}\right)^H \Omega_2 W_2^H \tilde{X}_{10} \\ &\stackrel{\text{def}}{=} \tilde{B}_1 + F. \end{aligned} \tag{3.5}$$

Consider the following two matrix pencils:

$$\tilde{X}_{10}^H A\tilde{X}_{10} - \lambda\tilde{X}_{10}^H B\tilde{X}_{10} \quad \text{and} \quad \tilde{A}_1 - \lambda\tilde{B}_1, \tag{3.6}$$

a relation for which is described by (3.4) and (3.5). First of all, the first pencil of (3.6) is in $\mathbb{D}(l)$, and moreover

$$c\left(\tilde{X}_{10}^H A\tilde{X}_{10}, \tilde{X}_{10}^H B\tilde{X}_{10}\right) \geqslant c(A, B). \tag{3.7}$$

We claim that under the assumption (3.1), $\tilde{A}_1 - \lambda \tilde{B}_1$ is in $\mathbb{D}(l)$, too. In fact, for every $x \in \mathbb{C}^l$, as noted by Li [5],

$$\begin{aligned}
&\sqrt{(x^H E x)^2 + (x^H F x)^2} \\
&\quad \leqslant \left\| \left((W_2^H \tilde{X}_{10})^H \Lambda_2 W_2^H \tilde{X}_{10}, (W_2^H \tilde{X}_{10})^H \Omega_2 W_2^H \tilde{X}_{10} \right) \right\|_2 \\
&\quad \leqslant \left\| (W_2^H \tilde{X}_{10})^H (\Lambda_2, \Omega_2) \begin{pmatrix} W_2^H \tilde{X}_{10} & \\ & W_2^H \tilde{X}_{10} \end{pmatrix} \right\|_2 \\
&\quad \leqslant \| W_2^H \tilde{X}_{10} \|_2^2 \qquad \left(\text{since } \alpha_j^2 + \beta_j^2 = 1\right) \\
&\quad \leqslant \left\| (W_2^H W_2)^{1/2} \right\|_2^2 \left\| (W_2^H W_2)^{-1/2} W_2^H \tilde{X}_{10} \right\|_2^2 \\
&\quad \leqslant \| X^{-1} \|_2^2 \left[\rho_2(\mathscr{X}_1, \tilde{\mathscr{X}}_1) \right]^2 \qquad [\text{see } (2.3)] \\
&\quad \leqslant \frac{\|(A, B)\|_2^2}{c(A, B)} \left[\rho_2(\mathscr{X}_1, \tilde{\mathscr{X}}_1) \right]^2 \qquad (\text{by Lemma 2.2}),
\end{aligned}$$

so

$$\begin{aligned}
&\sqrt{\frac{(x^H E x)^2 + (x^H F x)^2}{\left(x^H \tilde{X}_{10}^H A \tilde{X}_{10} x\right)^2 + \left(x^H \tilde{X}_{10}^H B \tilde{X}_{10} x\right)^2}} \\
&\quad \leqslant \frac{\|(A, B)\|_2^2}{c(A, B)} \cdot \frac{1}{c\left(\tilde{X}_{10}^H A \tilde{X}_{10}, \tilde{X}_{10}^H B \tilde{X}_{10}\right)} \left[\rho_2(\mathscr{X}_1, \tilde{\mathscr{X}}_1) \right]^2 \\
&\quad \leqslant \frac{\|(A, B)\|_2^2}{c(A, B)^2} \left[\rho_2(\mathscr{X}_1, \tilde{\mathscr{X}}_1) \right]^2 < 1 \qquad [\text{by } (3.1)].
\end{aligned}$$

By Lemma 2.3, we have proved $\tilde{A}_1 - \lambda \tilde{B}_1 \in \mathbb{D}(l)$, and moreover there is a permutation τ of $\{1, \dots, l\}$ such that

$$\max_{1 \leqslant j \leqslant l} \rho\Big((\alpha_j, \beta_j), \big(\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)} \big) \Big)$$

$$\leqslant \max_{\|x\|_2 = 1} \rho\Big(\big(x^H \tilde{X}_{10}^H A \tilde{X}_{10} x, x^H \tilde{X}_{10}^H B \tilde{X}_{10} x \big), \big(x^H \tilde{A}_1 x, x^H \tilde{B}_1 x \big) \Big)$$

$$= \max_{\|x\|_2 = 1} \frac{\| x^H E x x^H \tilde{B}_1 x - x^H F x x^H \tilde{A}_1 x |}{\sqrt{\big(x^H \tilde{X}_{10}^H A \tilde{X}_{10} x \big)^2 + \big(x^H \tilde{X}_{10}^H B \tilde{X}_{10} x \big)^2} \times \sqrt{\big(x^H \tilde{A}_1 x \big)^2 + \big(x^H \tilde{B}_1 x \big)^2}} \leqslant \frac{1}{c(A, B)}$$

$$\times \left| \max_{\|x\|_2 = 1} x^H \big(W_2^H \tilde{X}_{10} \big)^H \left(\frac{\Lambda_2 \tilde{B}_1 x - \Omega_2 x^H \tilde{A}_1 x}{\sqrt{\big(x^H \tilde{X}_{10}^H A \tilde{X}_{10} x \big)^2 + \big(x^H \tilde{X}_{10}^H B \tilde{X}_{10} x \big)^2}} \right) \big(W_2^H \tilde{X}_{10} \big) x \right|$$

$$\leqslant \frac{\| W_2^H X_{10} \|_2^2}{c(A, B)} \max_{\substack{\|x\|_2 = 1 \\ l+1 \leqslant j \leqslant n}} \rho\Big(\big(x^H \tilde{A}_1 x, x^H \tilde{B}_1 x \big), (\alpha_j, \beta_j) \Big). \tag{3.8}$$

With the help of the minimax property of the eigenvalue of definite pencils [8], we can prove

$$\max_{\substack{\|x\|_2 = 1 \\ l+1 \leqslant j \leqslant n}} \rho\Big(\big(x^H \tilde{A}_1 x, x^H \tilde{B}_1 x \big), (\alpha_j, \beta_j) \Big) \leqslant \max_{\substack{1 \leqslant i \leqslant l \\ l+1 \leqslant j \leqslant n}} \rho\big((\alpha_i, \beta_i), (\alpha_j, \beta_j) \big).$$

So it follows from (3.8) that

$$\max_{1 \leqslant j \leqslant l} \rho\Big((\alpha_j, \beta_j), \big(\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)} \big) \Big)$$

$$\leqslant \frac{\| W_2^H \tilde{X}_{10} \|_2^2}{c(A, B)} \max_{\substack{1 \leqslant i \leqslant l \\ l+1 \leqslant j \leqslant n}} \rho\big((\alpha_i, \beta_i), (\alpha_j, \beta_j) \big)$$

$$\leqslant \left(\frac{\| (A, B) \|_2}{c(A, B)} \right)^2 \max_{\substack{1 \leqslant i \leqslant l \\ l+1 \leqslant j \leqslant n}} \rho\big((\alpha_i, \beta_i), (\alpha_j, \beta_j) \big) \big[\rho_2 \big(\mathscr{X}_1, \tilde{\mathscr{X}}_1 \big) \big]^2.$$

Our proof is completed. ■

The following theorem deals with the Frobenius norm.

THEOREM 3.2. *Under the conditions of Theorem* 3.1, *there is a permutation* ω *of* $\{1,\dots,l\}$ *such that*

$$\sqrt{\sum_{j=1}^{l}\left[\rho\left((\alpha_j,\beta_j),\left(\tilde{\alpha}_{\omega(j)},\tilde{\beta}_{\omega(j)}\right)\right)\right]^2}$$

$$\leqslant \max_{\substack{1\leqslant i\leqslant l\\ l+1\leqslant j\leqslant n}} \rho\left((\alpha_i,\beta_i),(\alpha_j,\beta_j)\right)\frac{\|(A,B)\|_2}{c(A,B)}\cdot\frac{\delta}{\sqrt{1-\eta^2}}, \tag{3.9}$$

where $\delta = (\|(A,B)\|_2/c(A,B))^2\rho_2(\mathscr{X}_1,\tilde{\mathscr{X}}_1)\rho F(\mathscr{X}_1,\tilde{\mathscr{X}}_1)$, *and* η *is defined by* (3.1).

Proof. We have shown that under the conditions of Theorem 3.1, the second pencil in (3.6) is in $\mathbb{D}(l)$ and hence $W_1^H\tilde{X}_{10}$ is invertible. Using (3.4) and (3.5), we get

$$\left(\tilde{X}_{10}^H\right)A\tilde{X}_{10}\left(W_1^H\tilde{X}_{10}\right)^{-1}\Omega_1 - \left(\tilde{X}_{10}^H B\tilde{X}_{10}\right)\left(W_1^H\tilde{X}_{10}\right)^{-1}\Lambda_1$$

$$= \left(W_2^H\tilde{X}_{10}\right)^H(\Lambda_2 M\Omega_1 - \Omega_2 M\Lambda_1), \tag{3.10}$$

where $M = (W_2^H\tilde{X}_{10})(W_1^H\tilde{X}_{10})^{-1}$. By Lemma 2.1, there is $\tilde{P}_1\in\mathbb{C}^{l\times l}$ nonsingular such that

$$\tilde{P}_1^H\left(\tilde{X}_{10}^H A\tilde{X}_{10}\right)\tilde{P}_1 = \tilde{\Lambda}_1 \equiv \operatorname{diag}(\tilde{\alpha}_1,\dots,\tilde{\alpha}_l),$$

$$\tilde{P}_1^H\left(\tilde{X}_{10}^H B\tilde{X}_{10}\right)\tilde{P}_1 = \tilde{\Omega}_1 \equiv \operatorname{diag}\left(\tilde{\beta}_1,\dots,\tilde{\beta}_l\right),$$

where $\tilde{\alpha}_j^2+\tilde{\beta}_j^2 = 1$ for $i = 1,\dots,l$. Substituting those decompositions into (3.10), we get

$$\tilde{\Lambda}_1\left[\tilde{P}_1^{-1}\left(W_1^H\tilde{X}_{10}\right)^{-1}\right]\Omega_1 - \tilde{\Omega}_1\left[\tilde{P}_1^{-1}\left(W_1^H\tilde{X}_{10}\right)^{-1}\right]\Lambda_1$$

$$= \tilde{P}_1^H\left(W_2^H\tilde{X}_{10}\right)^H(\Lambda_2 M\Omega_1 - \Omega_2 M\Lambda_1). \tag{3.11}$$

It was proved by Li [3] that there exists a permutation ω of $\{1,\dots,l\}$ such that

$$\sqrt{\sum_{j=1}^{l}\left[\rho\left((\alpha_j,\beta_j),\left(\tilde{\alpha}_{\omega(j)},\tilde{\beta}_{\omega(j)}\right)\right)\right]^2}$$
$$\leqslant \left\|\left(W_1^H\tilde{X}_{10}\right)\tilde{P}_1\right\|_2$$
$$\times\left\|\tilde{\Lambda}_1\left[\tilde{P}_1^{-1}\left(W_1^H\tilde{X}_{10}\right)^{-1}\right]\Omega_1-\tilde{\Omega}_1\left[\tilde{P}_1^{-1}\left(W_1^H\tilde{X}_{10}\right)^{-1}\right]\Lambda_1\right\|_F. \tag{3.12}$$

Note that

$$\Lambda_2 M\Omega_1-\Omega_2 M\Lambda_1=G\circ M,$$

where $\circ$ denotes the Hadamard product of matrices and $G=(\alpha_{l+i}\beta_j-\beta_{l+i}\alpha_j)\in\mathbb{C}^{(n-l)\times l}$. Thus from (3.11) and (3.12), it follows that

$$\sqrt{\sum_{j=1}^{l}\left[\rho\left((\alpha_j,\beta_j),\left(\tilde{\alpha}_{\omega(j)},\tilde{\beta}_{\omega(j)}\right)\right)\right]^2}$$
$$\leqslant \max_{\substack{1\leqslant i\leqslant l\\ l+1\leqslant j\leqslant n}}\rho\left((\alpha_i,\beta_i),(\alpha_j,\beta_j)\right)\left\|\left(W_1^H\tilde{X}_{10}\right)\tilde{P}_1\right\|_2\|\tilde{P}_1^H\|_2$$
$$\times\left\|\left(W_2^H\tilde{X}_{10}\right)^H\right\|_2\left\|W_2^H\tilde{X}_{10}\left(W_1^H\tilde{X}_{10}\right)^{-1}\right\|_F$$
$$\leqslant \max_{1\leqslant i\leqslant l}\rho\left((\alpha_i,\beta_i),(\alpha_j,\beta_j)\right)\|W_1^H\|_2\left\|\left(W_1^H\tilde{X}_{10}\right)^{-1}\right\|_2\|\tilde{P}_1\|_2^2$$
$$\times\left\|\left(W_2^HW_2\right)^{1/2}\right\|_2^2\left\|\left(W_2^HW_2\right)^{-1/2}W_2^H\tilde{X}_{10}\right\|_2\left\|\left(W_2^HW_2\right)^{-1/2}W_2^H\tilde{X}_{10}\right\|_F. \tag{3.13}$$

We now bound factors in (3.13) one by one. By Lemma 2.2,

$$\|W_1^H\|_2\leqslant\left\|\begin{pmatrix}W_1^H\\W_2^H\end{pmatrix}\right\|_2=\|X^{-1}\|_2\leqslant\frac{\|(A,B)\|_2}{\sqrt{c(A,B)}},$$

$$\|\tilde{P}_1\|_2^2\leqslant\frac{1}{c\left(\tilde{X}_{10}^HA\tilde{X}_{10},\tilde{X}_{10}^HB\tilde{X}_{10}\right)}\leqslant\frac{1}{c(A,B)},$$

$$\left\|\left(W_2^HW_2\right)^{1/2}\right\|_2^2=\|W_2\|_2^2\leqslant\|X^{-1}\|_2^2\leqslant\frac{\|(A,B)\|_2^2}{c(A,B)},$$

and by (2.3) $\|(W_2^H W_2)^{-1/2} W_2^H \tilde{X}_{10}\|_p = \rho_p(\mathscr{X}_1, \tilde{\mathscr{X}}_1)$ for $p = 2$ or F. Finally, we deal with the hardest one, $\|(W_1^H \tilde{X}_{10})^{-1}\|_2$:

$$\begin{aligned}
&\tilde{X}_{10}^H W_1 W_1^H \tilde{X}_{10}\\
&\quad = \tilde{X}_{10}^H X^{-H} X^{-1} \tilde{X}_{10} - \tilde{X}_{10}^H W_2 W_2^H \tilde{X}_{10}\\
&\quad \geqslant \left(\frac{1}{\|X\|_2^2} - \|\tilde{X}_{10}^H W_2\|_2^2 \right) I\\
&\quad \geqslant \frac{1}{\|X\|_2^2} \left(1 - \|X\|_2^2 \left\| \tilde{X}_{10}^H W_2 (W_2^H W_2)^{-1/2} \right\|_2^2 \left\| (W_2^H W_2)^{1/2} \right\|_2^2 \right) I,
\end{aligned}$$

which produces

$$\begin{aligned}
\left\| \left(W_1^H \tilde{X}_{10} \right)^{-1} \right\|_2 &= \left\| \left(\tilde{X}_{10}^H W_1 W_1^H \tilde{X}_{10} \right)^{-1} \right\|_2^{1/2}\\
&\leqslant \frac{\|X\|_2}{\sqrt{1 - \|X\|_2^2 \left\| \tilde{X}_{10}^H W_2 (W_2^H W_2)^{-1/2} \right\|_2^2 \|X^{-1}\|_2^2}}\\
&\leqslant \frac{1}{\sqrt{c(A, B)}} \frac{1}{\sqrt{1 - \eta^2}}.
\end{aligned}$$

Substituting all those estimates into (3.13) leads to (3.9). ∎

For a general unitarily invariant norm, we have Theorem 3.3 below. To say that a norm $|||\cdot|||$ is *unitarily invariant* on $\mathbb{C}^{m\times n}$ means it satisfies, besides the usual properties of any norm, also

(1) $|||UAV||| = |||A|||$ for all $U \in \mathbb{U}_m$ and $V \in \mathbb{U}_n$;
(2) $|||A||| = \|A\|_2$ for any $A \in \mathbb{C}^{m\times n}$, rank $A = 1$.

The spectral norm $\|\cdot\|_2$ and the Frobenius norm $\|\cdot\|_F$ are two frequently used ones.

THEOREM 3.3. *Under the conditions of Theorem* 3.1, *there is a permutation σ of* $\{1, \dots, l\}$ *such that*

$$\begin{aligned}
&\left|\left|\left| \operatorname{diag}\left(\rho\left((\alpha_1, \beta_1), \left(\tilde{\alpha}_{\sigma(1)}, \tilde{\beta}_{\sigma(1)} \right) \right), \dots, \rho\left((\alpha_l, \beta_l), \left(\tilde{\alpha}_{\sigma(l)}, \tilde{\beta}_{\sigma(l)} \right) \right) \right) \right|\right|\right|\\
&\qquad \leqslant \pi \frac{\|(A, B)\|_2}{c(A, B)} \frac{\delta_u}{\sqrt{1 - \eta^2}}, \qquad (3.14)
\end{aligned}$$

where $\delta_u = (\|(A, B)\|_2 / c(A, B))^2 \rho_2(\mathscr{X}_1, \tilde{\mathscr{X}}_1) \||\sin \Theta(X_1, \tilde{X}_1)\||$, *and* η *is defined by* (3.1).

Proof. The proof is much like that of Theorem 3.2, except that the inequality (3.12) is replaced by the following one proved in Li [4]:

$$\begin{aligned}
&\left\|\left|\operatorname{diag}\left(\rho\left((\alpha_1, \beta_1), \left(\tilde{\alpha}_{\sigma(1)}, \tilde{\beta}_{\sigma(1)}\right)\right), \dots, \rho\left((\alpha_l, \beta_l), \left(\tilde{\alpha}_{\sigma(l)}, \tilde{\beta}_{\sigma(l)}\right)\right)\right)\right|\right\| \\
&\quad \leqslant \frac{\pi}{2} \left\|\left(W_1^H \tilde{X}_{10}\right) \tilde{P}_1\right\|_2 \\
&\quad\quad \times \left\|\left| \tilde{\Lambda}_1 \left[\tilde{P}_1^{-1}\left(W_1^H \tilde{X}_{10}\right)^{-1}\right] \Omega_1 - \tilde{\Omega}_1 \left[\tilde{P}_1^{-1}\left(W_1^H \tilde{X}_{10}\right)^{-1}\right] \Lambda_1 \right|\right\|
\end{aligned}$$

for some permutation σ of $\{1, \dots, l\}$. ∎

With the inequality (3.2) for the spectral norm in mind, we would find that the inequality (3.9) might not be the one as expected. Most probably, we would expect

$$\sqrt{\sum_{j=1}^{l} \left[\rho\left((\alpha_j, \beta_j), \left(\tilde{\alpha}_{\omega(j)}, \tilde{\beta}_{\omega(j)}\right)\right)\right]^2} \leqslant \max_{\substack{1 \leqslant i \leqslant l \\ l+1 \leqslant j \leqslant n}} \rho\left((\alpha_i, \beta_i), (\alpha_j, \beta_j)\right) \cdot \delta \tag{3.15}$$

to be true, which is stronger than (3.9). However, I was unable to prove it. A similar argument applies to (3.14).

4. CONCLUSION

We have proved three inequalities regarding the eigenvalues of Rayleigh quotient matrix pencils of a definite pencils. They are generalizations of their counterparts for Rayleigh quotient matrices proved in Li [2], Liu and Xu [6], and Sun [12]. Our result for the spectral norm improves Sun's [11], and our results for the Frobenius norm and general unitarily invariant norms are new, and seem to be improvable. We conjecture that (3.15) might be true.

The author wishes to thank referees for many helpful suggestions.

REFERENCES

1 S.-P. Han, Relative eigenvalues of Hermitian matrices, *Linear Algebra Appl.* 81:75–88 (1986).
2 R.-C. Li, On eigenvalues of a Rayleigh quotient matrix, *Linear Algebra Appl.* 169:249–255 (1992).
3 R.-C. Li, On perturbations of matrix pencils with real spectra, *Math. Comp.*, to appear.
4 R.-C. Li, On perturbations of matrix pencils with real spectra. II, in preparation.
5 R.-C. Li, A perturbation bound for definite pencils, *Linear Algebra Appl.* 179:191–202 (1993).
6 Xin-guo Liu and Ya-ge Xu, On the Rayleigh quotient matrix (in Chinese), *Math. Numer. Sinica* 12(2):208–213 (1990).
7 B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N.J., 1980.
8 G. W. Stewart, Perturbation bounds for the definite generalized eigenvalue problem, *Linear Algebra Appl.* 23:69–83 (1979).
9 G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic, New York, 1990.
10 J.-G. Sun, A note on Stewart's theorem for definite matrix pairs, *Linear Algebra Appl.* 48:331–339 (1982).
11 J.-G. Sun, Relative eigenvalues of a definite matrix pair, *Linear Algebra Appl.* 139:253–267 (1990).
12 J.-G. Sun, Eigenvalues of Rayleigh quotient matrices, *Numer. Math.* 59:603–614 (1991).
13 J.-G. Sun, Rayleigh quotient and residual of a definite pair, *J. Comput. Math.* 9:247–255 (1991).

Received 15 March 1993; final manuscript accepted 30 October 1993